

観光地検索システムの説明性向上のための既訪問スポットと検索結果の 対応付けの詳細化

潘 健太[†] 北山 大輔[†]

[†] 工学院大学大学院工学研究科情報学専攻 〒163-8677 東京都新宿区西新宿 1-24-2

E-mail: [†]tem18011@ns.kogakuin.ac.jp, ^{††}kitayama@cc.kogakuin.ac.jp

あらまし 近年、ユーザは Web 上の観光情報を活用して旅行計画を立てることが多くなっている。しかし、旅行は一般的に訪れたことがないスポットに行くことが多いため、観光情報を適切に理解することは困難である。そこで、訪問したことがある観光スポットの特徴を用いて、未訪問エリアの観光スポットを対応付けして、ユーザの未知なスポットに対する理解を支援することを提案する。本研究では、まず、観光スポットのユーザレビューを用いてレビューベクトルを作成する。次に、ユーザにとっての観光スポットの特徴を抽出するために、既訪問スポットレビューベクトルを用いて階層的クラスタリングすることによって、特徴クラスタを算出する。また、特徴クラスタに対応する検索結果スポットレビューを用いて類似度を計算し、既訪問スポットと検索結果スポットを関連付けを行い、関連性を説明するためのキーワードを抽出する。最後に、観光スポット間の対応関係を地図上に可視化する手法を用いて、プロトタイプシステムを構築し、既訪問スポットと検索結果スポットの説明性向上のための観光スポットの対応関係を評価する実験を行う。

キーワード 観光スポット, 理解支援, ユーザレビュー, 階層的クラスタリング, 可視化

1 はじめに

旅行先を決定するとき、旅行者は観光スポット検索サイトや観光情報に関連する書籍を見て観光スポットを選び、旅行計画を立てる。しかし、ユーザにとって訪問したいエリアは、訪問したことがなく不慣れであることが多い。そのため、エリア内に数多く存在する観光スポットから、自身の要求に合う観光スポットを見つけることは容易ではない。このとき、ユーザは観光スポット検索サイトのランキングやおすすめ情報を見て観光スポットを決めることが多くなると考えられる。ユーザの意識決定の材料の 1 つとして、Tripadvisor¹やじゃらん²などの観光スポット検索サイトがある。これらのサイトには特定の観光スポットを訪問したことのあるユーザがレビューを投稿し、観光スポットに関する豊富な情報が存在している。しかし、ユーザは検索エリアに関する事前知識がないため、どのスポットのレビューを読むべきか効率的に判断することは困難である。そのため、さまざまな観光スポットを効果的に理解するためには、ユーザが訪問した経験のあるスポットを使って検索結果スポットと比較することは効果的と考えられる。この考え方は、ユーザが以前に経験した物事を現在の物事に適用する一種の類推である [1]。類推は創造的思考に貢献すると指摘されてた [2]。類推に関する研究の多くは、ベースとなる学習データとターゲットとなる問題が与えられ、物事の特徴を問題の特徴にマッピングして問題を解決するもの [3] である。たとえば、日本に初めて訪れるフランス人旅行者に対し、検索結果スポットである東

京の「表参道」をパリにおける「シャンゼリゼ通り」と表現すると理解しやすいと考えられる。我々は、すでに既訪問スポットと検索結果スポットの対応関係を抽出する手法について取り込んでいる [4]。本稿では、既訪問スポットを用いてよりスポットの詳細化を取り込んでいる。本手法では、ユーザはすでに訪問した経験のあるスポットと未訪問エリアを入力する。そのとき、観光スポットのユーザレビューを用いて特徴ベクトルを作成する。次に、ユーザにとっての観光スポットの特徴を抽出するために、既訪問スポットレビューベクトルを用いて階層的クラスタリングすることで、特徴クラスタを算出する。また、特徴クラスタに対応する検索結果スポットレビューを用いて類似度を計算し、既訪問スポットと検索結果スポットを関連付けを行い、関連性を説明するためのキーワードを抽出する。最後に、観光スポット間の対応関係を地図上に可視化する手法 [5] を用いて、プロトタイプシステムを構築し、既訪問スポットと検索結果スポットの説明性向上のための観光スポットの対応関係を評価する実験を行う。抽出したキーワードを提示することで、ユーザの検索結果スポットに対する理解の支援を目指す。図 2 は提案手法の概念図である。

本論文の構成は下記のとおりである。2 節では、関連研究について述べる。3 節では、提案手法の概要について述べる。4 節では、検索結果スポットと既訪問スポットの対応関係による地図上の可視化手法について述べる。5 節では、クラスタの特徴語抽出に関する評価と考察について述べる。最後に、6 節では、まとめと今後の課題について述べる。

1 : <https://www.tripadvisor.com/>

2 : <https://www.jalan.net/kankou/>

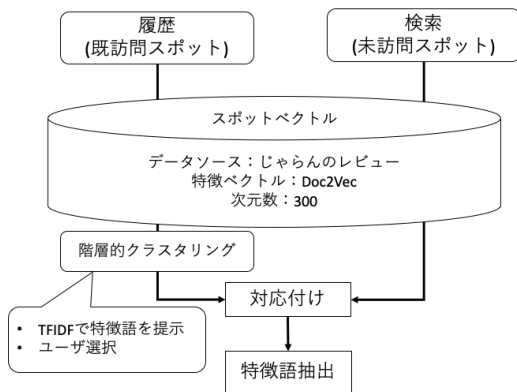


図 1 提案手法の概念図

2 関連研究

2.1 観光スポット検索と推薦システム

ユーザの体験履歴を利用した検索や推薦システムに関する研究は数多く発表されている。倉島ら [6] は、Flickr に投稿された写真のジオタグ情報を人々の旅行履歴として利用した旅行ルート推薦手法を提案した。この手法では、ユーザの現在地から行きやすい場所とユーザの興味に合致した場所に移動しやすいと仮定し、行動モデルを生成している。ユーザのジオタグ付き写真集合は、時間情報でソートすると個人の旅行履歴とみなすことができると考え、ジオタグ情報を利用してユーザの行動モデルを生成している。Kitamura ら [7] は、一般的な物体認識を用いて、過去の個人旅行写真から推定したユーザの旅行の嗜好に基づき観光地を推薦する方法を提案した。物体認識システムを用いて、写真で撮った被写体情報のキーワードを取得し、グラフ視覚化技術によってキーワードの共起を表現した。また、グラフの視覚化技術に基づいて旅行写真付きのグラフを視覚化するユーザインターフェイスを紹介した。

2.2 地理的情報の可視化と意味的な関係の可視化

本研究では、地理的情報の可視化と意味的な関係の可視化に取り組む。地理的な情報の可視化としては、抽出した情報を地図上にマッピングするのが一般的であり、本研究もそれに従っている。代表的な研究を以下に紹介する。櫻川ら [8] は、ソーシャルメディア上にアップロードされた写真のジオタグ情報と撮影時刻に基づいて写真の撮影者を分類する手法を提案した。分類された撮影者（観光者、在住者）ごとにホットスポットを可視化した。また、ジオタグ情報と撮影時刻以外にテキストタグを加えて、写真が撮影された地域で行われるイベントの穴場スポットを発見し、地図上に提示する手法を提案した [9]。

意味的な関係性の可視化としては、グラフモデル上でオブジェクト間のつながりを関連度で表現することが多い。上村ら [10] は、ユーザが投稿したタグ付き画像を用いたファッションスタイルの関係性の可視化手法を提案した。ファッションスタイルは類似するタグを空間座標に固定することによって関係を表している。本研究では、スポット間の意味的な対応関係を

可視化するために、検索結果スポットは地図上で実在する座標に固定され、配置の自由があるのは既訪問スポットのみという制約がある。

3 検索結果スポットと既訪問スポットの対応付け

我々は、ユーザの既訪問スポットに基づく検索結果スポットの対応付けの詳細化を提案する。まず、ユーザがすでに訪問した複数の観光スポットと訪問したい観光スポットエリア情報を入力する。本手法では、観光スポットのユーザレビューを用いて既訪問スポットレビューベクトルを生成する。検索結果スポットも同様にエリア内の各スポットレビューベクトルを求める。次に、ユーザの観光スポット訪問履歴から各訪問スポットに付随する個々のレビューを全て取得し、階層的クラスタリングによってレビューをクラスタに分類する。各クラスタを正規化ジニ分散指標とスポット数で算出したスコアの上位 3 件のクラスタの特徴語をユーザに提示する。そのあと、未訪問エリア内のスポットレビューをユーザが選択されたクラスタに対して平均類似度によってクラスタに分類する。ユーザ選択クラスタに属する既訪問スポットと検索結果スポットのレビューベクトルの類似性によって既訪問スポットを検索結果スポットと関連付ける。最後に、その関係性を説明するためのキーワードを抽出する。

3.1 スポットレビューを用いたレビューベクトル生成

本稿では、2016 年 9 月末までのじゃらんから得られた 1,481,838 件のレビューデータを使用し、分散表現 [11] を用いてレビューベクトルの作成する。本研究では、分散表現を計算するために Python のライブラリである gensim³ を利用する。学習方法として、Distributed Bag-of-Words を利用して、観光スポットのレビューを 300 次元のベクトルで表現する。既訪問スポットや検索結果スポットの各レビューベクトルは、形態素解析器である MeCab [13] に辞書「mecab-ipadic-NEologd」⁴を用いて、分かち書き（原型）したレビューを利用して作成する。

3.2 既訪問スポットのクラスタ分類と特徴語抽出

本研究では、ユーザが訪れた観光スポットにはユーザの観光スポットに対する嗜好が含まれていると仮定する。しかし、観光スポットには多くの場合、単一の要素ではなく複数の要素が存在する。そこで、ユーザが訪れた観光スポットに付随する全てのレビューをクラスタリングすることにより、レビューを観光スポットの持つ特徴ごとに分類する。レビューの分類には階層的クラスタリングを利用する。クラスタ間の距離の測定には、レビューベクトルのコサイン類似度による群平均法を用いる。群平均法とは、2 つのクラスタに属する要素間の全ての組み合わせの距離を求め、その平均値をクラスタ間の距離とする手法である。クラスタ分けの閾値は、山田ら [12] ユーザの嗜好に基づく観光スポット説明文の個人化手法の実験結果によって

3 : <https://radimrehurek.com/gensim/models/doc2vec.html>

4 : <https://github.com/neologd/mecab-ipadic-neologd/>

0.65 とする。

階層的クラスタリングによって分類したクラスタに対し、どのクラスタがユーザの嗜好を表しているかをスコア付けする。本研究では、以下の条件を満たすクラスタをユーザの嗜好を表すクラスタと定義する。

- 各既訪問スポットのレビューを満遍なく含んでいる
- より多くのスポットのレビューで構成されている

本研究では、クラスタを構成する訪問スポットごとのレビュー数の割合を算出するため、ジニ分散指標を用いる。ジニ分散指標とは各クラスの分散の総和を表す指標であり、クラスのばらつきが小さいと値は大きくなり、ばらつきが大きいほど値は小さくなる。既訪問スポット数を N 、クラスタ c に所属するレビューのうち、スポット s のレビューの割合を $P(s | c)$ とすると、ジニ分散指標は式 1 で表せる。

$$gini(c) = 1 - \sum_{s=1}^N P(s | c)^2 \quad (1)$$

また、各スポットのレビュー数とそのスポットの全レビュー数との割合を求めることにより、レビュー数の正規化を行なったものを正規化ジニ分散指標と呼称し、これを利用する。スポット s の総レビュー数を R_s とすると、正規化ジニ分散指標における $P(s | c)$ は式 2 で表せる。

$$P(s | c) = \frac{\frac{r_s}{R_s}}{\sum_{j=1}^N \left(\frac{r_j}{R_j} \right)} \quad (2)$$

ジニ分散指標は各クラスの分散の総和のみで値が決まる関係上、クラス数までは考慮されない。我々はより多くの訪問済みスポットで構成されたクラスタは、ユーザの観光に対する嗜好を表現していると考えたため、クラスタの構成スポット数による重み付けを行う。そのとき、あるスポットに関するレビューが 1 つしかない場合でも、そのクラスタを構成するスポットとしてカウントしてしまう問題がある。そこで本研究では、クラスタを構成するあるスポットのレビュー数と、そのスポットの全レビュー数との割合が閾値以上のものを、クラスタを構成するスポットとする。上記に示した閾値を 1% 以上と決定した。

よって、クラスタ c の嗜好の興味度 $pref(t)$ は、あるスポットに関するレビュー数の割合が閾値以上のスポット数を s_c とすると、式 3 で表せる。

$$pref(c) = gini(c) \times \log s_c \quad (3)$$

ユーザに提示するクラスタの特徴語は、スコア付されたクラスタのレビューを用いて、TFIDF 法で特徴語と TFIDF 値を求める。各クラスタの上位 10 件までの単語を特徴語としてユーザに提示する。全クラスタ数を C 、クラスタ c ($1 \leq r \leq C$) における単語 w の出現回数を $tf_c(w)$ 、全クラス数における単語 w の出現回数を $tf(w) = \sum_{c=1}^C tf_c(w)$ 、 w が出現するクラスタ数を $df(w)$ 、クラスタ c の総単語数を m_c 、全クラス数における総単語数を $M = \sum_{c=1}^C m_c$ 、クラスタ c の長さで正規化した単語の出現回数を $TF_c(w) = tf_c(w)/m_c$ とする。

$$\text{単語 } w \text{ の特徴量} = TF_c(w) * \log \frac{C}{df(w)} \quad (4)$$

3.3 対応するスポットの決定と特徴語抽出

3.2 節で求めたクラスタの特徴語に対して、ユーザが 1 つ選択する。上位 3 件までのクラスタに所属するレビューと未訪問エリア内のスポットの各レビューの類似度を計算する。未訪問エリア内のスポットのレビューとクラスタに所属する各レビューの類似度の平均をクラスタに対する検索結果スポットのレビューの類似度とする。検索結果スポットのレビューの類似度がもっとも大きい、かつ 0.125 以上のクラスタに検索結果スポットのレビューを分類する。ユーザが選択されたクラスタに所属する既訪問スポットを使って、検索結果スポットを説明する。したがって、検索結果スポットと既訪問スポットによって計算された類似度に基づいて関連付けを行う。そのためには、選択されたクラスタに所属する各既訪問スポットレビュー集合と各検索結果スポットレビュー集合を使う。類似度計算には、コサイン尺度 (式 5) を用いる。

$$\cos(t_{R_{S_f}}, t_{R_{S_u}}) = \frac{t_{R_{S_f}} \cdot t_{R_{S_u}}}{|t_{R_{S_f}}| \times |t_{R_{S_u}}|} \quad (5)$$

S_f を既訪問エリア内のスポット集合とする。 $t_{R_{S_f}}$ は選択されたクラスに所属する既訪問スポットのスポットのレビュー集合である。 S_u を未訪問エリア内のスポット集合とする。 $t_{R_{S_u}}$ は選択されたクラスに所属する検索結果スポットのスポットのレビュー集合である。

スポットの対応関係を示すだけでは、どのような点で対応するのかを理解するのは難しい。そこで、検索結果スポットと既訪問スポットの関係性を表すキーワードをユーザに提示する。すべてのレビューを 3.1 節と同様に形態素解析器 MeCab によって単語に分割する。ただし、助詞、助動詞、連体詞、記号、ストップワードを削除している。

キーワード抽出手順について説明する。まず、TFIDF 法を使って対象となる既訪問スポットと検索結果スポットの特徴語と TFIDF 値を求める。IDF 値を算出する集合は $S_{state} = \{s_1, s_2, \dots, s_n\}$ を用いる。

2 つのスポットに共通する特徴語の TFIDF 値の調和平均を用いて、対応付けしたスポットの説明可能なキーワードとして抽出する。各単語のスコアを式 6 によって定義する。

$$score(t, s_f, s_u) = \frac{2 \times tfidf(t, s_f) \times tfidf(t, s_u)}{tfidf(t, s_f) + tfidf(t, s_u)} \quad (6)$$

$tfidf(t, s_f)$ と $tfidf(t, s_u)$ は同じ単語に関する既訪問スポット s_f における TFIDF 値と検索結果スポット s_u における TFIDF 値を示している。単語スコアの上位 N 個の単語を説明情報としてユーザに提示する。

4 スポット間の対応関係を用いた可視化

3 で求めたある検索結果スポットに対して、入力した既訪問スポットがどのような関係にあるのか地図上で、一目で把握できるシステムが望ましい。そのため、我々は、観光スポット間の対応関係を地図上に可視化する手法 [5] を提案した。この手法では、ユーザは地図上で検索結果スポットを検索しているも

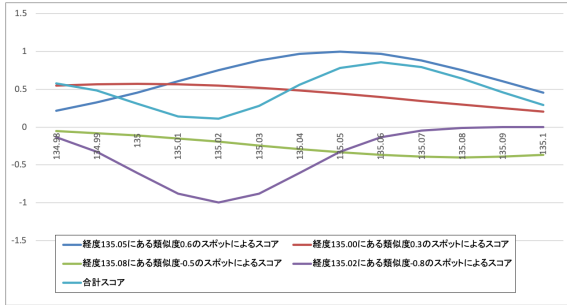


図 2 既訪問スポットの経度座標計算の例

のとし、まず、地図上の実際にスポットが存在する座標に検索結果スポットを配置する。次に、既訪問スポットの位置を決定する。これは検索結果スポットとの対応関係によって変化し、検索結果スポットとの類似度が高いと近く、類似度が低いと遠くなるように配置することで、意味的な位置関係を表現する。既訪問スポットの座標は、式 7 によって求める。

$$Score(s_f, T) = \sum_{s_u \in S_u} w(s_u, s_f) \times P_{s_u, s_f}(d(s_u, T)) \quad (7)$$

$w(s_u, s_f)$ は、ある既訪問スポット s_f と検索結果スポット s_u の類似度が正であれば 1.0、負であれば -0.5 を返す関数である。 T は、候補の座標である。 $d(s_u, T)$ は、ある座標 T から検索結果スポット s_u のユークリッド距離である。また、 P_{s_u, s_f} は、平均 μ が 0、標準偏差 σ が $(1 - |\cos(\text{vec}(s_u), \text{vec}(s_f))|) \times \alpha$ である正規分布を表す。 α は 1 である。 vec は検索結果スポットや既訪問スポットの特徴ベクトルを返す関数である。 $Score(s_f, T)$ がもっとも大きな座標 T を既訪問スポットの座標とする。図 2 は既訪問スポットの経度座標計算の例である。縦軸はスコアの値であり、横軸は経度である。4 つの検索結果スポットを用いて、 $Score(s_f, T)$ (合計スコア) を計算している。その結果、135.06 の座標がもっとも高くなる。緯度についても同様に行う。本稿では、近似として対象エリア内でランダムに 200 個の T を作り、それを用いた。

5 クラスタ分けに関する評価

5.1 実験内容

ユーザが訪問したことがある観光スポットから、ユーザの観光スポットに対する嗜好をクラスタする実験を行う。まず、既訪問スポットに対する嗜好を 3 つ想定したデータセットを、著者らが仮想的に 2 人分用意する。このデータセットを用いて既訪問スポットのレビューを取得し、階層的クラスタリングによってクラスタを作成しスコア付けを行う。そのあと、スコア上位のクラスタを構成するレビューを分析し、想定した嗜好を含むクラスタが作成されているかどうかを確認する。

本実験を利用するデータセットの内容を以下に示す。ユーザ A のデータセットを表 1、ユーザ B のデータセットを表 2 のように用意した。ユーザ A は「初詣」、「タワー」、「生物」を観光スポットに対する嗜好と仮定したユーザを想定している。ユーザ B も同様に、「遊園地」、「神社」、「山」を観光スポットに会

表 1 ユーザ A の既訪問スポット履歴

既訪問スポット	想定した嗜好	レビュー数
成田山新勝寺	初詣	787
川崎大師	初詣	241
東京スカイツリー	タワー	237
東京タワー大展望台	タワー	1875
海遊館	生物	2874
沖縄美ら海水族館	生物	3028
新宿御苑	その他	1147

表 2 ユーザ B の既訪問スポット履歴

既訪問スポット	想定した嗜好	レビュー数
富士急ハイランド	遊園地	1328
ユニバーサル・スタジオ・ジャパン	遊園地	760
高千穂神社	神社	135
天岩戸神社	神社	228
阿蘇山	山	444
富士山	山	761
新宿御苑	その他	1147

する嗜好とするユーザを想定した。また、ユーザの観光スポット訪問履歴には様々な嗜好が混じっていることから、「その他」としてジャンルの異なる観光スポットをそれぞれ 1 つずつ含めている。

5.2 実験結果と考察

表 3 と表 4 は正規化ジニ分散指標とスポット数による重み付けを用いたスコア付けの結果となっている。階層的クラスタリングによるクラスタ分けの結果を判断するため、スコア値が大きい上位 5 件を取りました。また、それぞれのクラスタ内の TFIDF 法で計算した TFIDF 値の上位の 10 件の単語を抽出して、クラスタがうまく作成されているかどうかを判断する要素として利用する。

ユーザ A のクラスタ分けの実験結果 (表 3) から、195 番のクラスタでは「初詣」に関する特徴語、134 番のクラスタでは「タワー」に関する特徴語、151 番のクラスタでは「生物」に関する特徴語が抽出されており、我々と仮定した嗜好に合うクラスタ分けができています。同様に、ユーザ B のクラスタ分けの実験結果 (表 4) から、131 番のクラスタでは「遊園地」に関する特徴語、80 番のクラスタでは「神社」に関する特徴語、56 番のクラスタでは「山」に関する特徴語が抽出されており、我々と仮定した嗜好に合うクラスタ分けができています。以上の結果から、クラスタの構成スポット数に夜重み付けが、ユーザの観光スポットに対する嗜好の抽出に有用であることがわかった。しかし、これによって、ユーザがスポット検索するとき、自分の嗜好に合うクラスタを選択することによって、より検索結果の対応付の詳細化することが可能と考えられる。

6 まとめと今後の課題

本研究では、ユーザが行きたい観光スポットが決まっていな

い場合に、ユーザの事前知識が不足しているため、観光検索サ

表 3 ユーザ A のクラスタ分け結果

クラスタ	スコア	クラスタ内の特徴語上位 10 件
151	0.57	水槽, ジンベイザメ, 水族館, 魚, ジンベエザメ, 迫力, サメ, 泳ぐ, 沖縄, イルカショー
134	0.56	東京タワー, 展望台, 夜景, スカイツリー, 東京, 景色, タワー, シンボル, 展望, 見える
195	0.14	成田, 成田山, うなぎ, お寺, 新勝寺, 成田山新勝寺, 初詣, 参道, 節分, 成田駅
155	0.07	記入, 館内, エスカレーター, 残念, 従業員, 悪さ, 対応, お天気, コーナー, 水族館
186	0.06	川崎大師, 渡す, お祓い, バッグ, 仲見世通り, 御籤, 一服, なんだか, APEC, すれ違う

表 4 ユーザ B のクラスタ分け結果

クラスタ	スコア	クラスタ内の特徴語上位 10 件
131	0.38	絶叫, アトラクション, 乗れる, トーマスランド, 乗り物, 乗る, ええじゃないか, ジェットコースター, ドドンパ, 遊園地
80	0.23	富士山, 火口, 合, 登る, 五合, 山頂, ガス, 阿蘇, 阿蘇山, 見える
56	0.14	天岩戸, 宮, 高千穂峡, 高千穂, 神社, 西本, 杉, 天安河原, 社務所, 天岩戸神社
53	0.08	参拝, 詣る, 参拝者, あまる, うんざり, ひとつだけ, 事情, 仏閣, 巫女, 手水舎
54	0.07	神話, 一回り, 学べる, 宮崎市内, 気象観測, 津, 美美, 駅そば, かけて, 説明

イトを使用してランキング, おすすめ情報やカテゴリなどを基に検索された観光スポットに対する理解が困難であることに着目した. 検索結果スポットに対する理解を支援するために, 検索結果スポットをユーザがすでに訪れたことがある既訪問スポットと比較することによって, 理解を支援する説明手法を提案した.

今後の課題としては, システムのインターフェースの作成および評価実験を行う予定である.

謝 辞

本研究の一部は, 平成 30 年度科研費基盤研究 (C)(課題番号: 18K11551) によるものです. ここに記して謝意を表すものとします.

文 献

- [1] D. Gentner, “Structure-Mapping: A Theoretical Framework for Analogy”, Cognitive Science, Vol.7, pp.155–170, 1983
- [2] K. J. Holyoak and P. Thagard, “Mental Leaps: Analogy in Creative Thought, MIT Press”, Journal of Japanese Society for Artificial Intelligence, Vol.11, No.3, pp.489, 1996
- [3] M. L. Gick and K. J. Holyoak, “Analogical Problem Solving”, Cognitive Psychology, Vol.12, pp.306–355, 1980
- [4] 潘 健太, 北山 大輔: ユーザの既訪問スポットの位置付けに基づく検索結果スポットの説明手法, 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2019), H7-2, 2019
- [5] 潘 健太, 北山 大輔: 地図上における検索結果スポットの説明性向上のための観光スポットの対応関係可視化手法, 観光情報学会 第 20 回研究会, 2019
- [6] T. Kurashima, T. Iwata, G. Irie and K. Fujimura., “Travel route recommendation using geotags in photo sharing sites”, CIKM ’10 Proceedings of the 19th ACM international conference on Information and knowledge management, pp.579-588, 2010
- [7] Kitamura, R. and Itoh, T.: Tourist Spot Recommendation Applying Generic Object Recognition with Travel Photos, In 2018 22nd International Conference Information Visualisation, pp.1-5, 2018
- [8] 櫻川 直洋, 廣田 雅春, 石川 博, 横山 昌平: ジオタグ付き写真の撮影者を在住者と観光者に分類することによるホットスポットの発見, 第 7 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2015), F6-3, 2015
- [9] 櫻川 直洋, 廣田 雅春, 石川 博, 横山 昌平: ジオタグ付き写真を用いたイベントとその穴場スポットの発見, 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2016), H5-3, 2016
- [10] 上村 幸汰, 桂井 麻里衣, 真木 勇人, 後藤 亮介: タグ付き画像を用いたファッションスタイルの関係性の可視化, 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2019), E8-2, 2019
- [11] Quoc V. Le and Tomas Mikolov, “Distributed representations of sentences and documents”, In Proceedings of the 31th International Conference on Machine Learning, ICML 2014, pp. 1188–1196, 2014
- [12] 山田 祥輝, 北山 大輔: ユーザの嗜好に基づく観光スポット説明文の個人化手法, 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2019), E3-1, 2019
- [13] T. Kudo, K. Yamamoto and Y. Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis”, Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237, 2004